Year: 2019

Dependence of flood peaks and volumes in modeled discharge time series: effect of different uncertainty sources

Brunner, Manuela I ; Sikorska-Senoner, Anna E

Abstract: Flood estimates needed for designing efficient and cost-effective flood protection structures are usually derived using observed peak discharges. This approach neglects, firstly, that floods are characterized not only by peak discharge but also by flood volume, and, secondly, that these characteristics are subject to modifications under climate and land use changes. Bivariate flood frequency analysis based on simulated discharge time series makes it possible to consider both flood peak and flood volume in design flood estimation. Further, this approach considers changes in discharge characteristics by using discharge series generated from climate time series used as an input for a hydrological model. Such series are usually not available at an hourly resolution but at a certain aggregation level (e.g. 24 h) and might not perfectly represent observed precipitation distributions. In this study, we therefore investigate how the aggregation and distribution of precipitation series and discharge distribution affect flood peaks and volumes and their dependence. We propose a framework for assessing the uncertainty in bivariate design flood estimates that is caused by different factors in the modeling chain, which consists of precipitation-discharge modeling, flood event sampling, and bivariate flood frequency analysis. The uncertainty sources addressed are precipitation aggregation and distribution, parameter and model uncertainty, and discharge resolution. Our results show that all of these uncertainty sources are relevant for design flood estimation and that the importance of the individual uncertainty sources is catchment dependent. Our results also demonstrate that substantial uncertainty is introduced already in the first step of the model chain because commonly used calibration procedures do not take into account the reproduction of flood volumes. Researchers should be aware of such deficiencies when performing bivariate flood frequency analysis on modeled discharge time series and should aim to tailor model calibration procedures to the problem at hand.

DOI: https://doi.org/10.1016/j.jhydrol.2019.03.024

Originally published at:
Brunner, Manuela I; Sikorska-Senoner, Anna E (2019). Dependence of flood peaks and volumes in modeled discharge time series: effect of different uncertainty sources. Journal of Hydrology, 572:620-629.

DOI: https://doi.org/10.1016/j.jhydrol.2019.03.024

Accepted Manuscript

Research papers

Dependence of flood peaks and volumes in modeled discharge time series: effect of different uncertainty sources

Manuela I. Brunner, Anna E. Sikorska

PII: S0022-1694(19)30264-1
DOI: https://doi.org/10.1016/j.jhydrol.2019.03.024
Reference: HYDROL 23558

To appear in: *Journal of Hydrology*

Received Date: 6 July 2018
Accepted Date: 12 March 2019

Please cite this article as: Brunner, M.I., Sikorska, A.E., Dependence of flood peaks and volumes in modeled discharge time series: effect of different uncertainty sources, *Journal of Hydrology* (2019), doi: https://doi.org/10.1016/j.jhydrol.2019.03.024

Dependence of flood peaks and volumes in modeled discharge time series: effect of different uncertainty sources

Manuela I. Brunner[1]

manuela.brunner@wsl.ch, Zürcherstrasse 111, CH-8903 Birmensdorf ZH

Anna E. Sikorska[2]

Winterthurerstrasse 190, CH-8057 Zürich

Abstract

Flood estimates needed for designing efficient and cost-effective flood protection structures are usually derived using observed peak discharges. This approach neglects, firstly, that floods are characterized not only by peak discharge but also by flood volume, and, secondly, that these characteristics are subject to modifications under climate and land use changes. Bivariate flood frequency analysis based on simulated discharge time series makes it possible to consider both flood peak and flood volume in design flood estimation. Further, this approach considers changes in discharge characteristics by using discharge series generated from climate time series used as an input for a hydrological model. Such series are usually not available at an hourly resolution but at a certain aggregation level (e.g. 24 h) and might not perfectly represent observed precipitation distributions. In this study, we therefore investigate how the aggregation and distribution of precipitation series and discharge distribution affect flood peaks and volumes and their dependence. We propose a framework for assessing the uncertainty in bivariate design flood estimates that is caused by different factors in the modeling chain, which consists of precipitation-discharge modeling, flood event sampling, and bivariate flood frequency analysis. The uncertainty sources addressed are precipitation aggregation and distribution, parameter and model uncertainty, and discharge resolution. Our results show that all of these uncertainty sources are relevant for design flood estimation and that the importance of the individual uncertainty sources is catchment dependent. Our results also demonstrate that substantial uncertainty is introduced already in the first step of the model chain because commonly used calibration procedures do not take into account the reproduction of flood volumes. Researchers should be aware of such deficiencies when performing bivariate flood frequency analysis on modeled discharge time series and should aim to tailor model calibration procedures to the problem at hand.

Keywords: `bivariate flood frequency analysis, data resolution, HBV, precipitation`

[1]Mountain Hydrology and Mass Movements, Swiss Federal Institute for Forest, Snow and Landscape research (WSL).
[2]Department of Geography, University of Zurich.

distribution, uncertainty assessment

1. Introduction

Flood estimates are needed for designing adequate flood protection structures, for improving flood management, and for increasing flood preparedness. Such estimates usually focus on one variable, namely, peak discharge. This narrow focus neglects that flood events are characterized by several variables including flood volume and duration. Flood peak and volume have been shown to be interdependent (Szolgay et al., 2015), and this dependence should be taken into account when estimating design floods (Mediero et al., 2010). Design flood estimation methods that consider both flood volume (V) and peak discharge (Q), as well as their dependence, rely on a bivariate analysis of these two variables and result in a set of Q–V pairs sharing the same frequency of occurrence (Serinaldi & Grimaldi, 2011; Brunner et al., 2017). Bivariate estimation procedures, like univariate procedures, usually focus on observed discharge time series that represent current climate and land use conditions. Both climate and land use are subject to changes that can impact flood magnitudes (Köplin et al., 2014; Archfield et al., 2016). Such changes need to be taken into account in flood estimation because flood protection structures are usually intended to be effective over a longer period of time. Changes in (bivariate) flood frequencies can be considered by estimating design floods using modeled discharge time series that account for changed climate and/or land use conditions. Such series can be derived using a hydrological model driven by meteorological time series generated by a climate model or by taking into account altered land use conditions.

These hydrological models are calibrated and validated against observations, and the resulting modeled series ideally represents the statistical characteristics of the observed time series, including not only the distribution of flood peaks and volumes but also their dependence. Various sources of uncertainty, however, may make it impossible to exactly reproduce these characteristics. These include uncertainties related to the hydrological model and its parameters and uncertainties resulting from the quality of the input data. These data are sometimes not available at a high resolution of 1 h and are more commonly available at a daily resolution, especially in the case of meteorological input generated by a climate model. One way of estimating hourly data from daily records is by distributing the daily totals to a certain number of hours during the day. Sikorska et al. (2018) showed that the disaggregation of daily precipitation totals into sub-daily precipitation input affects modeled peak discharges. However, it has not yet been investigated how precipitation disaggregation affects flood volume and the Q–V relationship. Besides precipitation duration and amount, the temporal precipitation distribution can also influence the discharge hydrograph and therefore peak discharge and flood volume. If the precipitation amount is not variable but equally/uniformly distributed over time, this results in smaller peak discharges (DVWK, 1984). It is not yet known whether a temporally variable precipitation distribution results in a different dependence between peak discharges and flood volumes (Q–V

dependence). Furthermore, it is not clear what happens to the Q–V dependence if we move from hourly to daily discharge data. The use of modeled time series instead of observed ones in bivariate flood frequency analysis therefore prompts new questions about the effect of various factors possibly influencing this Q–V dependence.

This study addresses a few of these questions and related sources of uncertainty. Generally, its aim is to assess the influence of various uncertainty sources on modeled Q–V relationships and bivariate frequency analysis. More specifically, we address the following questions:

- How does precipitation-discharge modeling affect the Q–V dependence and bivariate flood frequency analysis?
- How do temporal precipitation disaggregation and distribution affect Q–V dependence and bivariate frequency analysis?
- What is the effect of the temporal discharge resolution on Q–V dependence and bivariate frequency analysis?

To answer these questions, we propose a framework that enables an assessment of the influence of different factors in the modeling chain on the Q–V dependence and bivariate frequency analysis. The model chain consists of precipitation-discharge modeling, flood event sampling, and bivariate frequency analysis. The uncertainty is assessed by modifying a single step in the modeling chain, simulating a discharge time series, and re-estimating bivariate design quantiles; these estimated values are then compared to design quantiles obtained by analyzing either an observed discharge time series or a modeled reference time series (Montanari & Koutsoyiannis, 2012). We applied this approach to a set of nine study catchments in Switzerland to assess how well Q–V dependence can be reproduced under different data quality and model settings.

2. Material and methods

2.1. Study catchments and data

The dataset included the following catchments: Surb at Döttingen (C1), Wigger at Zofingen (C2), Mentue at Yvonand (C3), Kleine Emme at Littau (C4), Areuse at Boudry (C5), Muota at Ingenbohl (C6), Grosstalbach at Isenthal (C7), Kander at Hondrich (C8), and Lütschine at Gsteig (C9) (Figure 1) and was used in a previous study by Sikorska et al. (2018). The meso-scale catchments have areas between 40 and 500 km^2, have mean elevations between 511 and 2050 m a.s.l., and are not influenced by hydropower plants or by large lakes. For a more detailed description of the catchments, the reader is referred to Table 1 in Sikorska et al. (2018). The catchments can be assigned to one of three groups: 1) catchments located on the Swiss Plateau which are mainly dominated by precipitation (C1–C3), 2) catchments located in the Pre-Alps and

Jura Mountains (C4–C6), which are influenced by a mix of precipitation and snow melt processes, and 3) catchments in the Alps where snow pack and glaciers exert a dominant control over discharge (C7–C9). The floods in the first group are mainly characterized by short-rain floods. Floods in the second and third group are often influenced by rain or rain-falling-on snow, whereas those in the third group can be influenced by snow- and glacier-melt (Sikorska et al., 2015b).

The dataset consisted of observed discharge, precipitation, and temperature data. The observed discharge time series are available at an hourly resolution and span a period of 35.5 years on average. The areal precipitation and the mean catchment temperature time series were derived from hourly time series at gauging stations operated by MeteoSwiss (Federal Office of Meteorology and Climatology MeteoSwiss, 2003) using the Thiessen polygon method (Schumann, 1998).

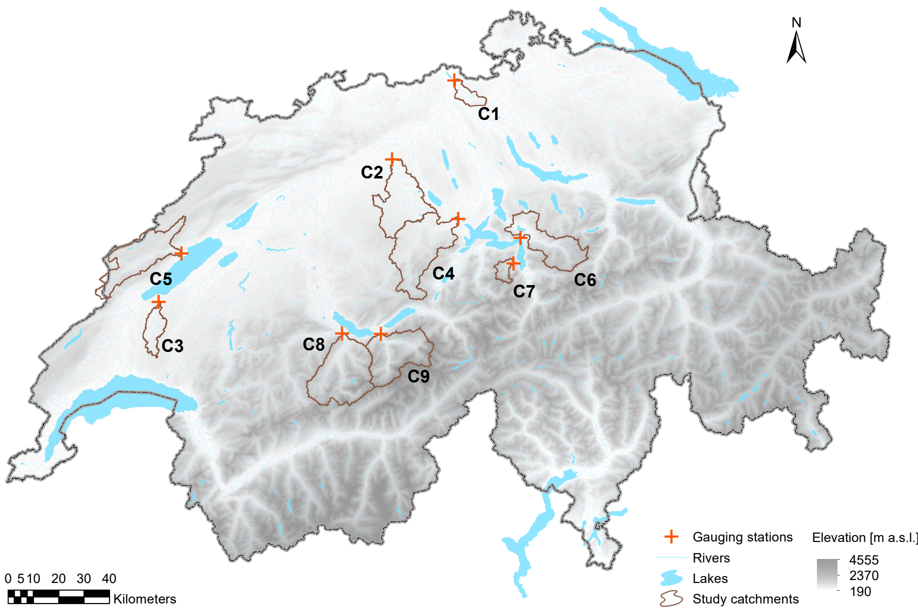


Figure 1: Map of the nine study catchments in Switzerland.

2.2. Assessment framework

The effect of various uncertainty sources on Q–V dependence and bivariate flood frequency was assessed within a model chain consisting of precipitation-discharge modeling, flood sampling, and bivariate flood frequency analysis, as illustrated in Figure 2. The uncertainty sources addressed included 1) the disaggregation level of the input precipitation data, 2) the temporal distribution of the input precipitation data, 3) precipitation-discharge model uncertainty, 4) the uncertainty of the parameters in the precipitation-discharge model, and 5) the resolution of the discharge data. The general principle was to change one element in the modeling chain a time and to run through all the steps in the modeling chain and assess the effect that this single change had on the bivariate distribution of Q and V expressed by their empirical copula. The empirical copula was used to estimate bivariate quantiles lying on an isoline of equi-probable events for a return period of $T = 10$. We focused on this return period because, as a return period of 100 years, it is of interest in engineering applications and estimates are reliable even with a relatively small flood sample of roughly 30 events, as was available in this study. The isolines generated using discharge time series resulting from multiple data and model configurations were used as input for the uncertainty analysis. This process consisted of comparing these isolines to isolines obtained by either observed discharge time series or a modeled reference time series. The different uncertainty sources and the procedure for assessing their effect on Q–V dependence are described in more detail in the following sections.

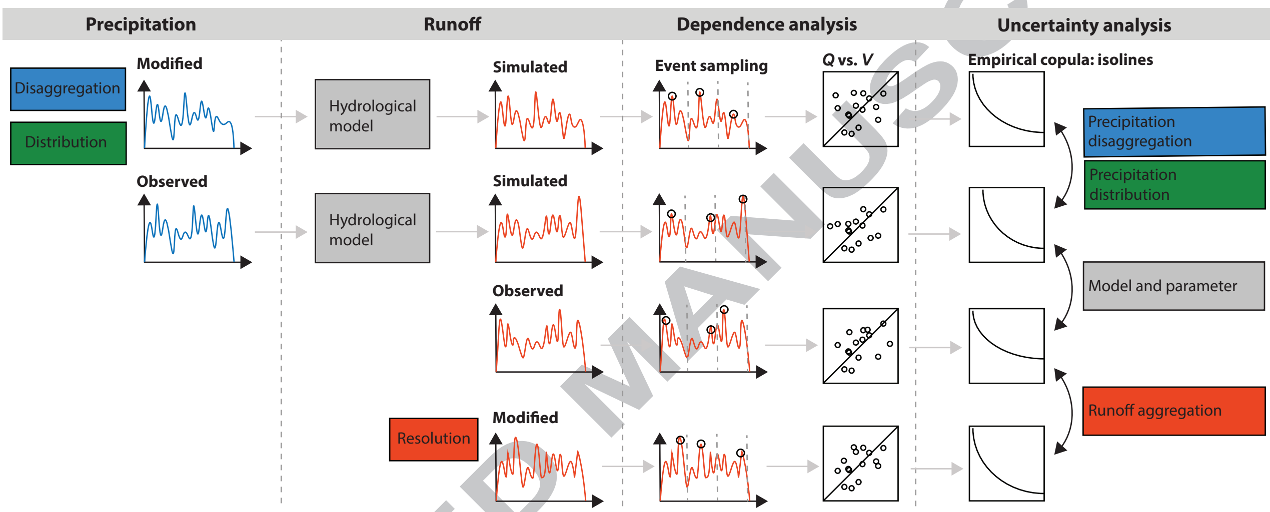


Figure 2: Illustration of the modeling setup and the uncertainty assessment framework applied in this study.

2.3. Precipitation disaggregation levels and distribution

We assessed the effects of temporal precipitation disaggregation as well as distribution on the dependence between Q and V. This was done by modifying observed precipitation time series and feeding them into a hydrological model to generate discharge time series showing the effect of changes in the precipitation time series on flood behavior. The effect of temporal precipitation disaggregation was assessed by aggregating the observed hourly data to daily totals and then uniformly redistributing them to m consecutive hours using 1, 2, 3, 6, 12, and 24 h as values for m. The first hour of precipitation was chosen randomly for each day. These newly disaggregated precipitation series were fed into the hydrological model and the effect on the Q–V dependence was assessed by comparing simulation results to the observed hourly data. For a more detailed description of this aggregation and disaggregation procedure, the reader is referred to Sikorska et al. (2018). The effect of temporal precipitation distribution on Q–V dependence was compared to the uniform distributions described above by redistributing observed precipitation according to an intensity evolution pattern found to generate high peak discharges by (DVWK, 1984) and therefore called the *worst case* distribution. To generate this distribution, 20% of the total precipitation amount was assigned to the first 30% of the event duration, 50% was assigned to the following 20% of the duration, and 30% was assigned to the remaining 50% of the duration.

2.4. Hydrological modeling

Discharge time series simulated with a hydrological model were compared to observed discharge time series at an hourly resolution to determine how well hydrological models reproduce the Q–V dependence in the observed data. The simulated time series were generated by driving the HBV model in the version 'light' (Lindström et al., 1997; Seibert & Vis, 2012) with observed and modified precipitation series. The model has five major modeling routines: precipitation excess and snow melt processes, soil moisture, groundwater and response, glacier melt, and discharge routing in the river.

The model parameters were calibrated with the observation period 1990–1999 and validated with the period 2000–2009 using the observed precipitation time series. The genetic algorithm and Powell optimization were used within a multi-objective framework to identify suitable parameter sets (Seibert, 2000). In the pre-calibration runs, we tested different objective functions and their compositions with different weights assigned. Among these, the Kling–Gupta efficiency function (Gupta et al., 2009), which is commonly applied for hydrological model calibration, led to the best overall model performance. Thus, this objective function was used for the final calibrations, which consisted of 2000 runs for the genetic algorithm and 100 runs for the local Powell's optimization. A clear advantage of using the Kling–Gupta efficiency (R_{KG}) is the possibility of its decomposition into three components, i.e., correlation (r), bias ratio (β) and variance (α) that support the interpretation of the model performance. The achieved median Kling–Gupta efficiency over all nine catchments was 0.89 in the calibration period (with $r = 0.89$, $\beta = 0.99$ and $\alpha = 1.00$) and 0.85 in the validation period ($r = 0.89$, $\beta = 0.99$ and $\alpha = 0.97$). All three components were close to the desired value of 1 with a slightly worse performance for the correlation component (r), which represents the agreement in temporal dynamics between observed and simulated discharge time series. The other two components indicated that the variability of simulated and observed discharge was very similar (α) and that the bias ratio (β) in simulated volumes was negligibly small (see Table 1 for the model performance in individual catchments). These results indicated a good overall model performance and thus permitted its application for the purpose of this study.

For the assessment of parameter uncertainty, the model calibration was repeated in each catchment 100 times, which resulted in 100 optimized parameter sets of a similar model performance. A number of 100 parameter sets was chosen as a compromise. On the one hand, it allowed for the representation of model parameter uncertainty and on the other hand, it was still computationally feasible. The same model set-up (including model parameters) was used for each of the modified precipitation time series for the simulation

Table 1: Kling–Gupta efficiency and its components for the calibration and validation period for the best parameter set in the nine study catchments.

Catchment	Calibration				Validation			
	R_{KG}	r	β	α	R_{KG}	r	β	α
C1	0.88	0.89	1.00	1.00	0.81	0.83	1.10	0.97
C2	0.89	0.89	1.00	1.00	0.70	0.85	1.20	1.20
C3	0.79	0.84	0.87	1.00	0.85	0.86	0.99	1.00
C4	0.88	0.88	1.00	0.99	0.87	0.89	0.99	0.94
C5	0.91	0.92	0.98	1.00	0.86	0.89	0.95	0.94
C6	0.90	0.91	0.99	1.00	0.84	0.90	1.00	1.10
C7	0.75	0.76	0.92	0.99	0.50	0.75	1.10	1.40
C8	0.91	0.92	0.98	0.96	0.88	0.91	0.98	0.92
C9	0.89	0.90	1.00	0.95	0.87	0.89	0.97	0.93

period 1980-2014.

2.5. Discharge resolution

The effect of selecting a certain discharge resolution was assessed by comparing the Q–V dependence in the observed data to the Q–V dependence in discharge time series with a coarser temporal resolution generated using a simple aggregation scheme. The hourly precipitation data were aggregated to values once every $m = 2$, 3, 6, 12, or 24 h. Note that the aggregation of precipitation data was based on precipitation totals while keeping an hourly resolution of the data. Thus, the hydrological model was still run at an hourly resolution but with aggregated precipitation data.

2.6. Event sampling

The dependence analysis relied on flood events sampled from the observed or simulated discharge time series described in the previous paragraphs. The flood events were sampled using an annual maxima strategy with respect to peak discharges, following a procedure similar to that used by Brunner et al. (2017) for peak-over-threshold events. An event window of 120 h (starting 48 h before peak discharge and ending 72 h after peak discharge) was defined to delimit the corresponding flood event. A window of 120 h was chosen because visual inspection showed that event durations did not exceed this duration. Within this window, the start of the event was defined as the point when the hydrograph exceeded 0.3 times the peak discharge for the first time. The end of the event was analogously defined as the point when the hydrograph fell below 0.3 times the peak discharge for the first time. The factor 0.3 was considered suitable for identifying events that were meaningful in terms of the hydrographs selected. The volume of the event was determined by integrating the area under the hydrograph during the event identified. The event sampling resulted in 35 events per simulation. We did not complete a season- or flood-type-specific analysis, which would have been desirable from a process perspective (Brunner et al., 2018, 2017), because subdividing the limited sample into individual seasons or flood types would have reduced the sample size too much for a reliable statistical analysis. The peak discharge–flood volume pairs were subsequently used to determine Q–V dependence and to perform bivariate flood frequency analysis.

2.7. Dependence analysis

We used Kendall's tau as a dependence measure to define rank correlation (Kendall, 1937). Furthermore, we computed the empirical copula of the Q–V pairs as an indicator of their bivariate distribution (Grimaldi et al., 2016). Similar to the empirical distribution, the empirical copula can, according to Joe (2014) be defined for multivariate independent and identically distributed (i.i.d.) data after a transformation to ranks. It can be expressed as

$$\widehat{C}_n(u) = n^{-1}\sum_{i=1}^{n} I([r_{ij} - \frac{1}{2}]/n \leq u_j, j = 1, ..., d), \tag{1}$$

where $r_{1j}; ...; r_{nj}$ are the (increasing) ranks for the j_{th} variable, u is a uniformly distributed variable, d is the dimension of the copula, I is the indicator function, and n is the number of observations. The ranks are adjusted to the uniform score $[r_{ij} - \frac{1}{2}]/n$. This empirical copula puts a mass of n^{-1} at the pairs $([r_{i1} - \frac{1}{2}]/n, ..., [r_{id} - \frac{1}{2}]/n)$ for $i = 1, ..., n$. The empirical copula was used to estimate bivariate design quantiles for a return period of 10 years. In contrast to univariate frequency analysis, where the definition of a return period is straightforward, bivariate frequency analysis requires the selection of a bivariate probability distribution to be used for the calculation of return periods (Brunner et al., 2016). We used the joint OR return period, which relies on the assumption that both peak discharge and flood volume are equally critical in the generation of a flood hazard. It can be expressed as follows:

$$T(u, v) = \frac{\mu}{1 - C(u, v)}, \tag{2}$$

where u and v correspond to the ranks adjusted to the uniform score, C corresponds to the copula (the empirical copula in this context), and μ corresponds to the mean inter-arrival time between events (1 in this context because annual maxima were used). All pairs (u, v) that are at the same probability level of the copula, have the same bivariate return period, and lie on an isoline. Isolines computed from flood samples resulting from observed and simulated discharge time series served as the input for the uncertainty analysis.

2.8. Uncertainty analysis

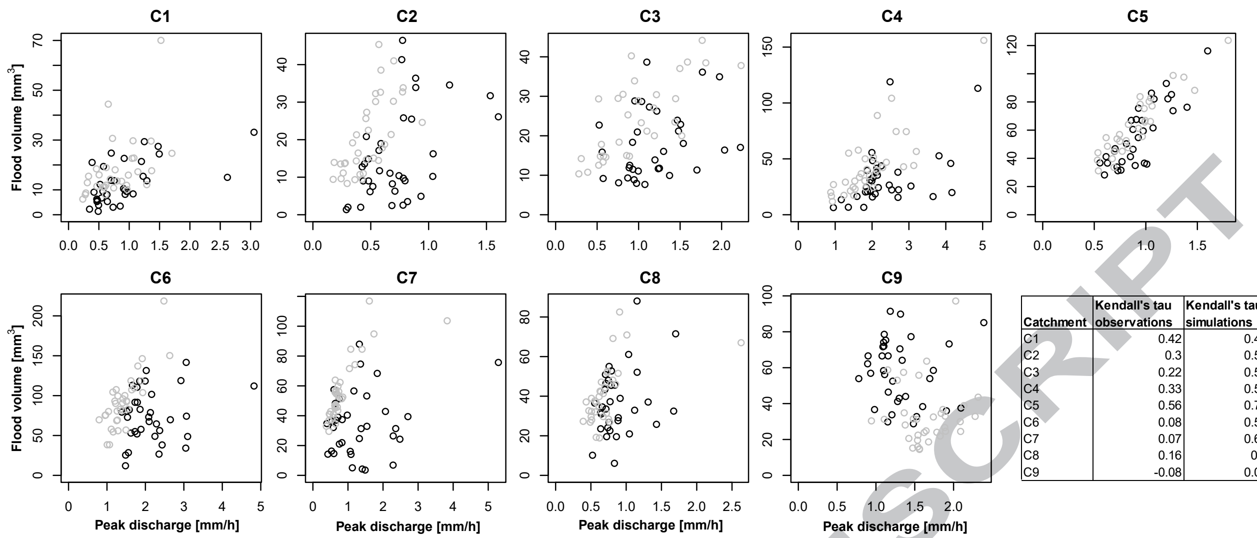
The uncertainty in the bivariate distributions was assessed by comparing isolines derived from Q–V pairs corresponding to a flood event with a return period of $T = 10$ years derived from a reference time series (either observed discharge, or discharge simulated with observed precipitation) to isolines derived from discharge time series obtained after perturbing the model chain (either precipitation or discharge time series). The uncertainty of the isolines was quantified by determining the area between the reference isoline and the perturbed isoline or by determining the mean area between several isolines and the reference isoline. As mentioned above, we assessed the effect of five uncertainty sources on bivariate flood frequencies:

1. **Disaggregation level of the input precipitation data**. The isolines obtained by simulations using modified precipitation inputs at different disaggregation levels using the uniform distribution were compared to the isoline obtained by the discharge simulations based on observed precipitation data. In this way, the effect of precipitation input data disaggregated from daily totals on the Q–V dependence was assessed.
2. **Temporal distribution of the input precipitation data**. The isoline obtained by the discharge simulations generated with the worst case precipitation distribution was compared to the isoline relying on the discharge time series generated using the uniform precipitation distribution.
3. **Combined model and parameter uncertainty**. The combined model and parameter uncertainty was assessed by comparing the isolines obtained from the observed flood events with the isoline derived from the flood events that were identified within discharge time series simulated using observed precipitation time series. The model uncertainty alone could not be assessed (see Sect. 4.4).
4. **Model parameter uncertainty**. Isolines obtained by discharge time series generated using 100 parameter sets and the observed precipitation time series were compared to the isoline obtained by discharge time series generated using the best parameter set and the observed precipitation time series.
5. **Resolution of the discharge data**. The isolines of discharge time series at five aggregation levels were compared to the isoline of the observed discharge time series.

3. Results

3.1. Reproduction of Q–V dependence by the hydrological model

Generally, Q–V dependence was higher for low-elevation than high elevation catchments (Table in Figure 3). The Q–V pairs extracted from the simulated series (eight out of nine catchments) showed a higher dependence than the Q–V pairs extracted from the observed discharge time series (Figure 3). This Q–V dependence overestimation was rather weak in catchment C5, where the Q–V dependence was relatively high already in the observations, and in C9, but it was stronger in the remaining catchments except for C1, where the simulated dependence was very similar to the observed dependence. Generally, simulated flood volumes were higher than observed volumes. Peak discharges were underestimated in high-elevation catchments.



Catchment	Kendall's tau observations	Kendall's tau simulations
C1	0.42	0.41
C2	0.3	0.57
C3	0.22	0.56
C4	0.33	0.59
C5	0.56	0.72
C6	0.08	0.56
C7	0.07	0.68
C8	0.16	0.5
C9	-0.08	0.07

Figure 3: Peak discharge–flood volume pairs for the flood events, extracted from the discharge observations (black dots) and the discharge simulations with observed precipitation time series (gray dots). Each catchment (C1 to C9) is shown in a separate panel and the scales of the axes vary among the nine plots.

3.2. Effect of precipitation disaggregation and distribution on Q–V dependence

The left two columns in Figure 4 show the empirical densities of flood volumes and peak discharges resulting from the 35 events extracted from the simulations generated using the five precipitation disaggregation levels (shades of blue) and the worst case precipitation distribution (green). The lower precipitation disaggregation levels (1–3 h) led to an over-representation of low volumes compared to high volumes in most catchments but C5–C7. The contrary was the case for peak discharges, where lower disaggregation levels led to an over-representation of high peak discharges and the densities were best represented by a resolution of roughly 6 h. The worst case precipitation distribution reproduced both the flood peak and volume densities well.

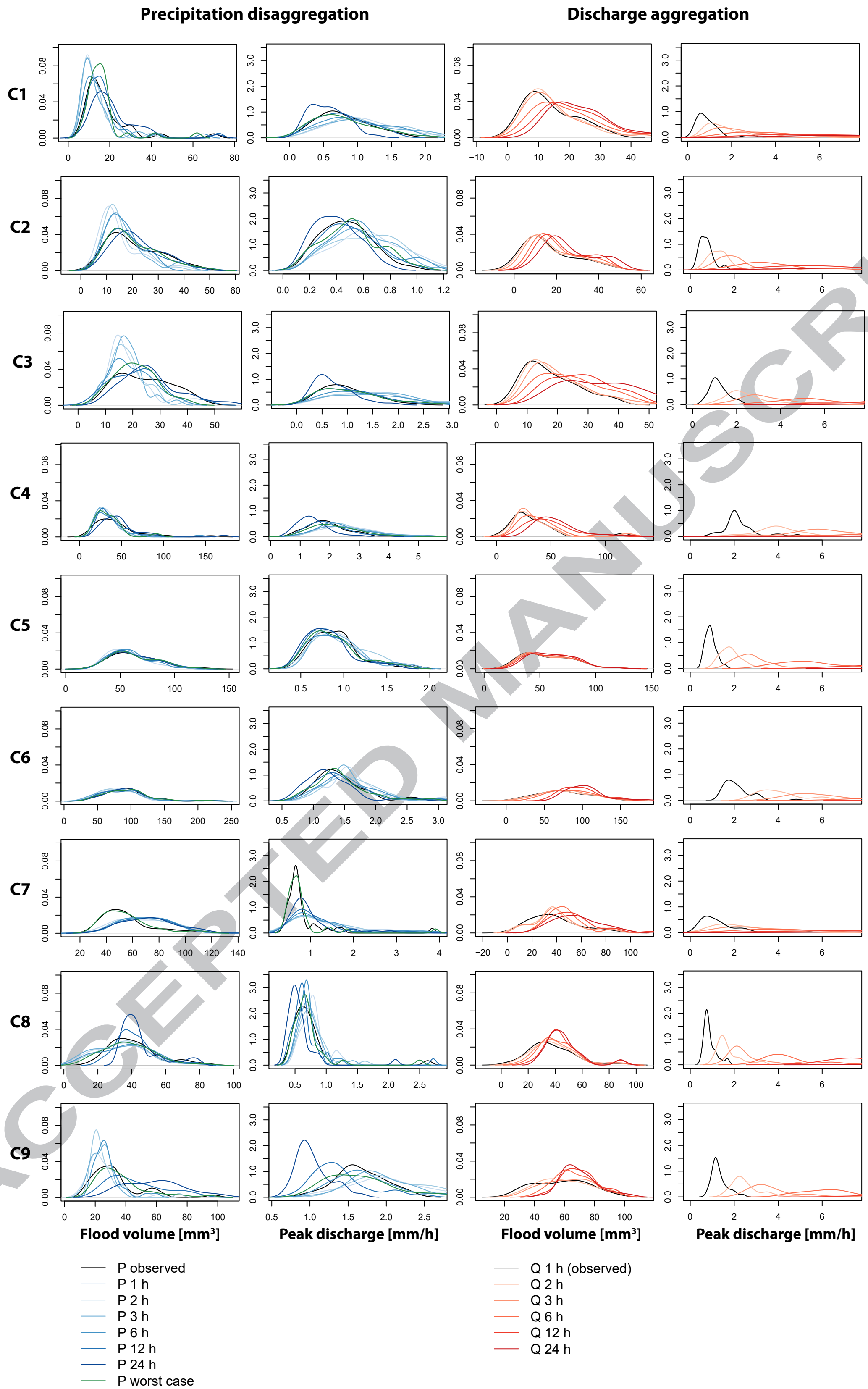


Figure 4: Empirical densities of peak discharges and flood volumes obtained for different precipitation disaggregation levels and the worst case distribution (left two columns) and for different temporal discharge resolutions (right two columns) for the nine study catchments (rows).

For some catchments, the dependence between peak discharge and flood volume was reproduced well independent of the precipitation disaggregation level chosen (C2, C5, C6, C7), while in others (C1, C3, C4, C8, C9) it was reproduced better with certain disaggregation levels (Figure 5). In these latter catchments (except for C8), the Q–V dependence was overestimated for the lower precipitation disaggregation levels (1 h–3 h) and better reproduced by the higher disaggregation levels (6 h–24 h). In C8, the dependence was also better reproduced by the higher disaggregation levels and underestimated by the lower disaggregation levels. The worst case scenario for precipitation (green bar) reproduced the Q–V dependence well and was closer to the Q–V dependence values obtained by the observed precipitation time series (black bar) than several of the uniform precipitation distributions represented by the different disaggregation levels.

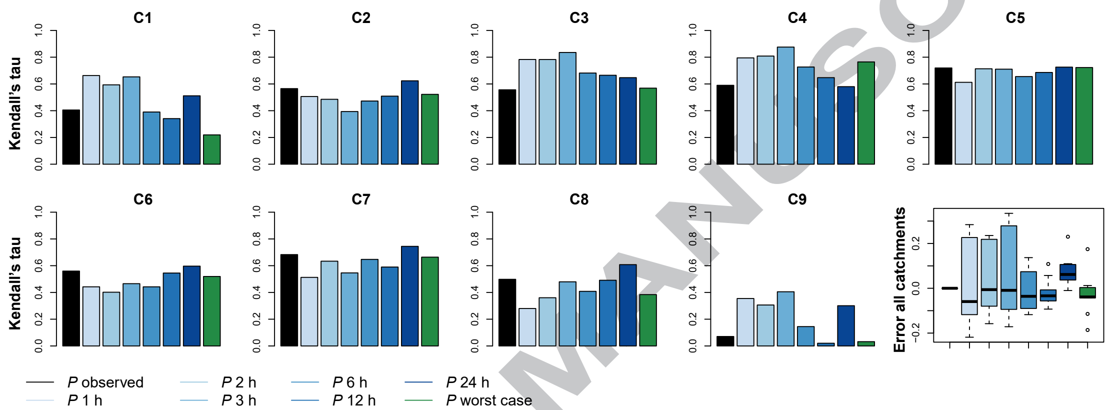


Figure 5: Kendall's tau of peak discharge and flood volume for different stations (black) and precipitation disaggregation levels (shades of blue) and for the worst case distribution (green). Errors, i.e., the difference between simulations and observations, over all catchments are shown in the boxplots in the lower right corner for all precipitation disaggregation levels.

3.3. Effect of discharge resolution on Q–V dependence

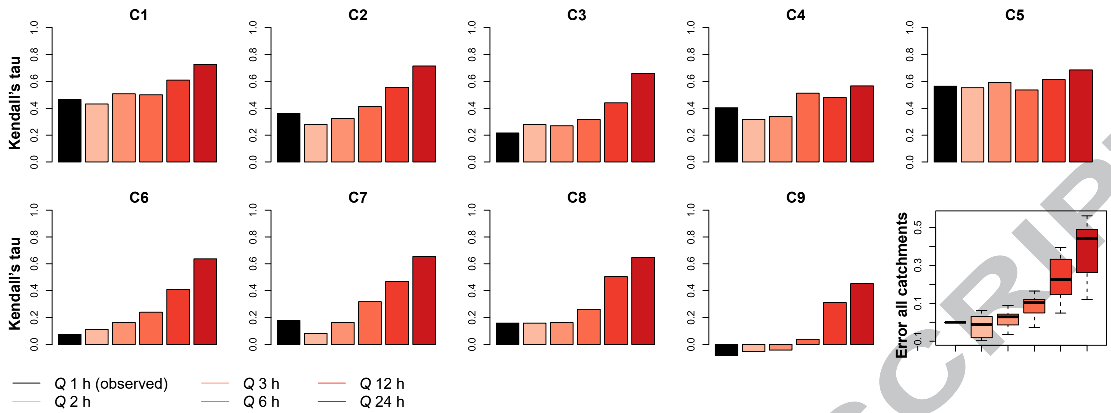


Figure 6: Q–V dependence in terms of Kendall's tau for the nine study catchments and different discharge resolution levels: 1 h (observations), 2 h, 3 h, 6 h, 12 h, and 24 h. Errors, i.e., the difference between simulations and observations, over all catchments are shown in the boxplots in the lower right corner for all discharge resolution levels.

The right two columns of Figure 4 show the densities of flood volumes and peak discharges referring to 35 events obtained by different temporal discharge resolutions each. A decrease in the temporal resolution led to an over-representation of high flood volumes and of high peak discharges. The Q–V dependence increased with decreasing discharge resolution (Figure 6) and was reproduced best for the high discharge resolutions (2 h–3 h). Q–V dependence was still reasonably well reproduced for a 6 h discharge resolution in most catchments. Lower temporal resolutions led to a strong increase in Q–V dependence.

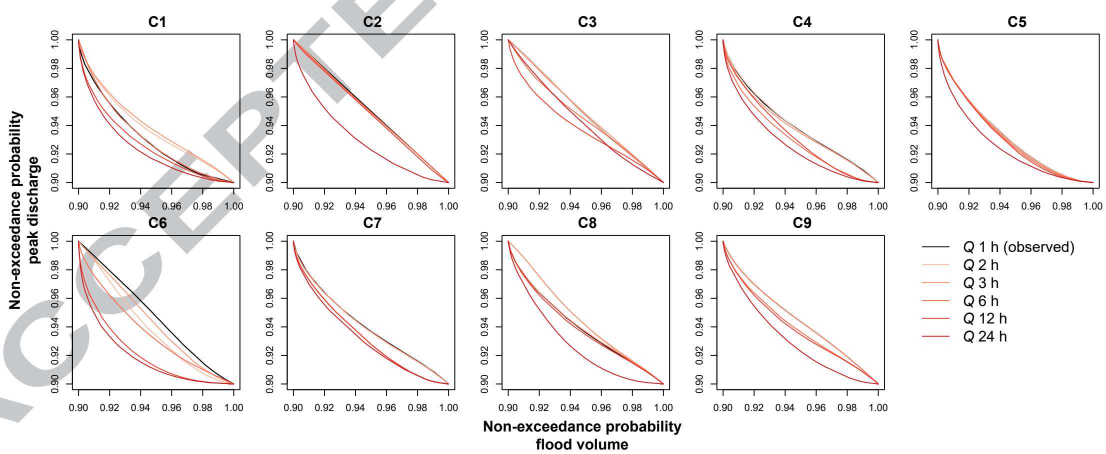


Figure 7: Isolines for an event with $T = 10$ for the nine study catchments and different temporal discharge resolutions: 1 h (observations), 2 h, 3 h, 6 h, 12 h, and 24 h.

Figure 7 shows isolines for an event with $T = 10$ derived from the empirical copulas resulting from different temporal discharge resolutions. Darker isolines represent lower temporal resolution levels. The lower the temporal resolution was, the lower the estimated joint quantiles were. The joint quantiles were reproduced best for the higher discharge resolutions as indicated by the similarity of the corresponding isolines.

3.4. Uncertainty analysis

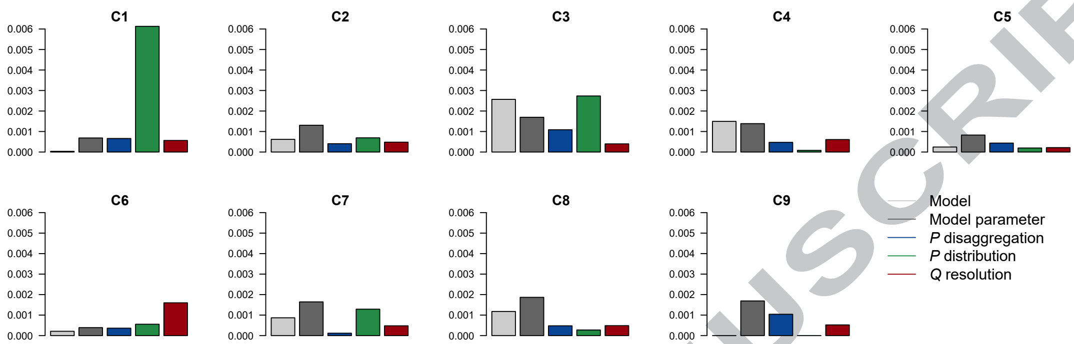


Figure 8: Uncertainty of the $T = 10$ isolines due to five uncertainty sources, expressed as the mean absolute error of the simulated isolines compared to a reference isoline, for the nine study catchments.

Figure 8 displays the uncertainty of estimated isolines for an event with $T = 10$ resulting from five uncertainty sources: model, model parameter, precipitation disaggregation, precipitation distribution, and discharge resolution. The relative importance of the different uncertainty sources varied from catchment to catchment. In C1 and C3, the largest uncertainty came from the precipitation distribution. In C2, C5, C7, C8, and C9, model parameter uncertainty was the most important source. In C6, the largest uncertainty came from the discharge resolution, and in C4 it came from the hydrological model. In addition, in C3, C4, C7, and C8, the uncertainty of the hydrological model regarding the bivariate design quantiles was also substantial. Although, their overall model performance in terms of the Kling–Gupta efficiency and its three components was not necessarily worse than in the remaining catchments (see Table 1). Therefore, the importance of the model uncertainty on bivariate quantiles cannot be linked to the overall model performance achieved in the catchment. Generally, the uncertainty coming from precipitation disaggregation and distribution and discharge resolution were in most catchments smaller than the uncertainty coming from the model parameters (see Sect. 4.4).

4. Discussion

4.1. Reproduction of Q–V dependence by the hydrological model

Our results show that using the hydrological model employed in this study to simulate hourly discharge time series leads to an overestimation of the Q–V dependence, which is in line with findings by Ben Aissia et al. (2014), who also found a higher Q–V dependence in simulated than in observed series. This overestimation was especially pronounced in catchments at higher elevations where the Q–V dependence is typically lower than in low–elevation catchments because they experience a mix of flood types (Gaál et al., 2015). This overestimation of the dependence can be explained by the temporal resolution chosen for modeling (hourly) and by the calibration procedure chosen. Convective events, which are short, intense, and locally restricted, typically show a very fast rising limb and are well visible at a temporal resolution of 1 h while such events are smoothed out when looking at daily data. The peak discharges become therefore more visible compared to the flood volumes in the hourly time series than in the daily series. This results in a lower Q–V dependence for hourly than daily series. This low dependence was not well reproduced by the hydrological model because the convective events with a high peak/volume ratio were poorly reproduced. The volume of convective events was overestimated in the modeled series compared to that for the observations. This led to an increase in the Q–V dependence by using the hydrological model. This overestimation would not be an issue if daily data were used because the peak/volume ratios at that scale are less severe.

Besides this resolution issue, the model calibration procedure chosen might have influenced the reproduction of Q–V dependence. In this study, we chose the Kling–Gupta efficiency (R_{KG}) for calibrating the hydrological model. Overall, a good model performance was achieved for all catchments as assessed by R_{KG} (0.89 and 0.85 in calibration and validation, respectively) and its three individual components, i.e., correlation (r), bias ratio (β) and variability (α), which were all close to the desired value of 1. This indicates that i) there was a good agreement between the simulated and observed discharge time series ($r = 0.89$), ii) discharge volumes were well fitted ($\beta = 0.99$), and iii) the variability of simulated and observed discharge time series was very similar ($\alpha \approx 1$) for both the calibration and the validation period. Note, however, that this calibration procedure focuses on the entire discharge time series and is not particularly adjusted to specific events (e.g., convective storms). Thus flood volume was, in contrast to peak discharge, not explicitly considered in the calibration procedure. The peak discharges were therefore possibly more reliably reproduced than the flood volumes. Still, the computation of the Q–V dependence relies on the volume of specific events. When interested in bivariate flood frequencies of flood peaks and volumes, one should therefore consider ways to integrate flood volume or even the Q–V dependence into the calibration procedure, e.g., by developing new objective functions tailored to the problem at hand, which would increase the robustness of the model (Pool et al., 2018). Pool et al. (2017) showed that the objective function selected for model calibration strongly influences the estimation accuracy of particular discharge signatures. We should be aware of a potential overestimation of the Q–V dependence when using modeled discharge time series for design flood estimation. These findings suggest that the choice of the hydrological model and its calibration procedure are important factors for bivariate design quantile estimation.

Our analysis showed that, besides model uncertainty, which was assessed only indirectly (see Sect. 4.4), parameter uncertainty also plays an important role in bivariate design quantile estimation. Parameter uncertainty has been found to be important in previous studies (e.g., Clark et al., 2016; Sikorska et al., 2015a) and should therefore be accounted for in design flood estimation studies using modeled discharge data.

4.2. Effect of precipitation disaggregation and distribution on Q–V dependence

Our results demonstrate that both precipitation disaggregation and precipitation distribution play a role in the reproduction of Q–V dependence. The Q–V dependence was over all catchments found to be reproduced the best at a disaggregation level of 6 or 12 h which agrees well with the results of Sikorska et al. (2018) who found that a disaggregation level of 6 h was sufficient to model peak discharges. A uniform distribution of precipitation either over– or underestimated Q–V dependence depending on the catchment. In contrast, the non-uniform worst case distribution at a 1 h resolution more accurately reproduced peak discharges, flood volumes, and the dependence structure in the observations in all catchments. This finding implies that not only the disaggregation level of the precipitation but also its distribution is important for reproducing flood volumes, peak discharges, and their dependence. This is in line with findings by Singh (1997), where the temporal distribution of rainfall was decisive in determining the magnitude of peak discharge. He found that peak discharges were higher for variable precipitation than for uniformly distributed precipitation, which is also the case in our study when comparing the non-uniform distribution to the uniform distributions at high disaggregation levels. In contrast, low disaggregation levels where the daily precipitation totals were distributed into 1–2 h led to higher peak discharges, because of high intensities, than the worst case distribution where the precipitation was divided into more time steps.

4.3. Effect of discharge resolution on Q–V dependence

Discharge resolution was found to be very important for reproducing Q–V dependence. If we work with daily data instead of hourly data, which is often the case in practice, we need to be aware that Q–V dependence is overestimated compared to when hourly observations are used. We have shown that working with discharge time series with a low temporal resolution leads to an underestimation of bivariate design quantiles. Such an underestimation might lead to a conservative design of flood retention basins or other flood protection structures and therefore have adverse effects on flood risk. It is very important that we keep this in mind when choosing appropriate discharge samples for bivariate design flood estimation.

4.4. Uncertainty analysis

Our results show that different sources of uncertainty are important in design flood estimation, depending on the catchment of interest. However, model parameter uncertainty was found to be one of the most important uncertainty sources in all the considered catchments. This source of uncertainty should therefore not be neglected. It has to be stressed, however, that parameter uncertainty, represented in this study by the 100 optimized parameter sets of the hydrological model, also comprises the uncertainty of the hydrological model itself and other uncertainty sources affecting hydrological modelling, i.e., the uncertainty of the observed precipitation and discharge time series used for model calibration (Renard et al., 2010; Sikorska & Renard, 2017). Although the model uncertainty was indirectly assessed separately, its direct contribution could not be assessed. The uncertainty due to uncertainty in the input time series was neither quantified. For this purpose a Bayesian uncertainty inference would be required (Gelman et al., 2013).

The other uncertainty sources in bivariate design quantile estimation should not be disregarded either, as they may affect the Q–V dependence. In some catchments, precipitation disaggregation was the most important source of uncertainty, while discharge resolution was a more important source in other catchments. Our results show that it is important to accurately represent the precipitation distribution and to work with highly resolved discharge data, if possible, to accurately represent the Q–V dependence in the observed or simulated time series.

4.5. Limitations and perspectives

This study was performed using a data set of nine Swiss catchments. This dataset represents different discharge regimes in Switzerland but was too small to detect a distinct pattern in the uncertainty behavior of catchments with respect to data resolution and precipitation-discharge modeling. The case study results are therefore not easily transferable to other catchments. The uncertainty assessment framework and the methods behind it can, however, be transferred to other catchments and regions. As mentioned above, the hydrological modeling procedure could be optimized with respect to the reproduction of flood volumes and Q–V dependence by taking these factors directly into account in model calibration. We did not investigate how the spatial distribution of precipitation affects flood volumes and modeled Q–V dependence as we focused on a lumped hydrological model. Spatial distribution has been found to be an important factor in simulating streamflow (Lobligeois et al., 2014) and in determining peak discharges (Singh, 1997). Its effect on flood volumes and peak discharges should be assessed by using a distributed instead of a lumped hydrological model. This study focused on a catchment-specific analysis without distinguishing between different flood types, which would have been preferable for the analysis of Q–V dependence, because of the limited sample size resulting from a subdivision of the flood samples into several flood types. One way to consider flood types would be to pool events of the same type within a region of similar catchments (Gaál et al., 2016).

The focus of this study has been on the reproduction of the variables peak discharge and flood volume and their dependence by using disaggregated precipitation data. Flood events are also characterized by a particular hydrograph shape. Useful procedures have recently been proposed for taking into account these shapes in design flood estimation (Brunner et al., 2016; Serinaldi & Grimaldi, 2011). Such procedures can also be applied in a climate change context using modeled discharge time series (Brunner et al., 2018). As a next step, it would therefore be important to look at the effect of the input precipitation disaggregation level and of the distribution not only on the magnitude of events in terms of peak discharges and flood volumes but also on hydrograph shapes, as proposed by Singh (1997).

5. Conclusions

Flood volumes and Q–V dependence are not necessarily well reproduced in modeled discharge time series because commonly used model calibration procedures do not explicitly take into account flood volumes in addition to peak discharges. It is crucial to be aware of this issue when using such simulated discharge time series for bivariate design flood estimation. Methods should therefore be developed to more accurately represent flood volumes and the Q–V dependence in hydrological models. Both the temporal disaggregation of precipitation and its distribution have an effect on peak discharges, flood volumes, and their dependence, and therefore on bivariate design flood estimation. This temporal distribution is especially important for reproducing observed peak discharges and flood volumes. The input data for modeling studies or the observed data should therefore be chosen with care with respect to their temporal resolution and distribution. A temporal resolution of 6 to 12 h is recommended for an accurate representation of Q–V dependence and the distribution must not be uniform but temporally variable. Not only the choice of precipitation input disaggregation but also the choice of the discharge resolution is important. Daily data generally lead to higher peak discharges and flood volumes than hourly discharge data and to a higher Q–V dependence. This higher dependence can lead to an underestimation of bivariate design quantiles compared to when using the actual dependence of the observed variables. We should be aware of the fact that the choice of a certain data set over another one might significantly influence the results of bivariate flood frequency analyses and the conclusions drawn from them.

Acknowledgements

The two authors acknowledge funding from the Swiss Federal Office for the environment (FOEN) via the contract 15.0003.PJ/Q292-5096 (granted to MB) and contract 15.0054.PJ/O503-1381 (granted to AS). discharge data were provided by the FOEN and the Canton of Argovia (Surb-Döttingen), and precipitation and temperature data were provided by MeteoSwiss. The authors thank the two anonymous reviewers for their useful comments.

References

Archfield, S. A., Hirsch, R. M., Viglione, A., & Blöschl, G. (2016). Fragmented patterns of flood change across the United States. *Geophysical Research Letters*, *43*, 232–239. doi:10.1002/2016GL070590.

Ben Aissia, M.-A., Chebana, F., Ouarda, T., Roy, L., Bruneau, P., & Barbet, M. (2014). Dependence evolution of hydrological characteristics, applied to floods in a climate change context in Quebec. *Journal of Hydrology*, *519*, 148–163. doi:10.1016/j.jhydrol.2014.06.042.

Brunner, M. I., Seibert, J., & Favre, A.-C. (2016). Bivariate return periods and their importance for flood peak and volume estimation. *Wire's Water*, *3*, 819–833. doi:10.1002/wat2.1173.

Brunner, M. I., Sikorska, A. E., & Seibert, J. (2018). Bivariate analysis of floods in climate impact assessments. *Science of the Total Environment*, *616-617*, 1392–1403. doi:10.1016/j.scitotenv.2017.10.176.

Brunner, M. I., Viviroli, D., Sikorska, A. E., Vannier, O., Favre, A.-C., & Seibert, J. (2017). Flood type specific construction of synthetic design hydrographs. *Water Resources Research*, *53*. doi:10.1002/2016WR019535.

Clark, M. P., Wilby, R. L., Gutmann, E. D., Vano, J. A., Gangopadhyay, S., Wood, A. W., Fowler, H. J., Prudhomme, C., Arnold, J. R., & Brekke, L. D. (2016). Characterizing Uncertainty of the Hydrologic Impacts of Climate Change. *Current Climate Change Reports*, *2*, 55–64. doi:10.1007/s40641-016-0034-x.

DVWK (1984). *Regeln zur Wasserwirtschaft. Arbeitsanleitung zur Anwendung von Niederschlag-Abfluss-Modellen in kleinen Einzugsgebieten. Teil II: Synthese*. Technical Report DVWK Hamburg und Berlin.

Federal Office of Meteorology and Climatology MeteoSwiss (2003). *SwissMetNet : The MeteoSwiss reference monitoring network*. Technical Report MeteoSwiss Zurich.

Gaál, L., Szolgay, J., Bacigál, T., Kohnová, S., Hlavčová, K., Výleta, R., Parajka, J., & Blöschl, G. (2016). Similarity of empirical copulas of flood peak-volume relationships: A regional case study of North-West Austria. *Contributions to Geophysics and Geodesy*, *46*, 155–178. doi:10.1515/congeo-2016-0011.

Gaál, L., Szolgay, J., Kohnová, S., Hlavčová, K., Parajka, J., Viglione, A., Merz, R., & Blöschl, G. (2015). Dependence between flood peaks and volumes: a case study on climate and hydrological controls. *Hydrological Sciences Journal*, *60*, 968–984. doi:10.1080/02626667.2014.951361.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. (3rd ed.). Boca Raton: CRC Press. Taylor & Francis Group.

Grimaldi, S., Petroselli, A., Salvadori, G., & De Michele, C. (2016). Catchment compatibility via copulas: A non-parametric study of the dependence structures of hydrological responses. *Advances in Water Resources*, *90*, 116–133. doi:10.1016/j.advwatres.2016.02.003.

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*, 80–91. doi:http://dx.doi.org/10.1016/j.jhydrol.2009.08.003.

Joe, H. (2014). *Dependence modeling with copulas*. Boca Raton: Chapman & Hall/CRC.

Kendall, M. G. (1937). A New Measure of Rank Correlation. *Biometrika*, *30*, 81. doi:10.2307/2332226.

Köplin, N., Schädler, B., Viviroli, D., & Weingartner, R. (2014). Seasonality and magnitude of floods in Switzerland under future climate change. *Hydrological Processes*, *28*, 2567–2578. doi:10.1002/hyp.9757.

Lindström, G., Johansson, B., Persson, M., Gardelin, M., & Bergström, S. (1997). Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology*, *201*, 272–288. doi:10.1016/S0022-1694(97)00041-3.

Lobligeois, F., Andréassian, V., Perrin, C., Tabary, P., & Loumagne, C. (2014). When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events. *Hydrology and Earth System Sciences*, *18*, 575–594. doi:10.5194/hess-18-575-2014.

Mediero, L., Jiménez-Alvarez, A., & Garrote, L. (2010). Design flood hydrographs from the relationship between flood peak and volume. *Hydrology and Earth System Sciences*, *14*, 2495–2505. doi:10.5194/hess-14-2495-2010.

Montanari, A., & Koutsoyiannis, D. (2012). A blueprint for process-based modeling of uncertain hydrological systems. *Water Resources Research*, *48*, 1–15. doi:10.1029/2011WR011412.

Pool, S., Vis, M., & Seibert, J. (2018). Evaluating model performance: a non-parametric variant of the Kling-Gupta efficiency. *Hydrological Sciences Journal*, *63*, 1941–1953. doi:10.1080/02626667.2018.1552002To.

Pool, S., Vis, M. J. P., Knight, R. R., & Seibert, J. (2017). Streamflow characteristics from modelled runoff time series - importance of calibration criteria selection. *Hydrology and Earth System Sciences*, *21*, 5443–5457. doi:10.5194/hess-21-5443-2017.

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., & Franks, S. W. (2010). Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. *Water Resources Research*, *46*, 1–22. doi:10.1029/2009WR008328.

Schumann, A. H. (1998). Thiessen polygon. In *Encyclopedia of Hydrology and Lakes* (pp. 648–649). Dordrecht: Springer Netherlands. doi:10.1007/1-4020-4497-6{_}220.

Seibert, J. (2000). Multi-criteria calibration of a conceptual runoff model using a genetic algorithm. *Hydrology and Earth System Sciences*, *4*, 215–224. doi:10.5194/hess-4-215-2000.

Seibert, J., & Vis, M. J. P. (2012). Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. *Hydrology and Earth System Sciences*, *16*, 3315–3325. doi:10.5194/hess-16-3315-2012.

Serinaldi, F., & Grimaldi, S. (2011). Synthetic design hydrographs based on distribution functions with finite support. *Journal of Hydrologic Engineering*, *16*, 434–446. doi:10.1061/(ASCE)HE.1943-5584.0000339.

Sikorska, A., & Renard, B. (2017). Calibrating a hydrological model in stage space to account for rating curve uncertainties: general framework and key challenges. *Advances in Water Resources*, *105*, 51–66. doi:10.1016/j.advwatres.2017.04.011.

Sikorska, A. E., Montanari, A., & Koutsoyiannis, D. (2015a). Estimating the uncertainty of hydrological predictions through data-driven resampling techniques. *Journal of Hydrologic Engineering*, *20*, A4014009. doi:10.1061/(ASCE)HE.1943-5584.

Sikorska, A. E., Viviroli, D., & Seibert, J. (2015b). Flood type classification in mountainous catchments using crisp and fuzzy decision trees. *Water Resources Research*, *51*, 7959–7976. doi:10.1002/2015WR017326.

Sikorska, A. E., Viviroli, D., & Seibert, J. (2018). Effective precipitation duration for runoff peaks based on catchment modelling. *Journal of Hydrology*, *556*, 510–522. doi:10.1016/j.jhydrol.2017.11.028.

Singh, V. P. (1997). Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph. *Hydrological processes*, *11*, 1649–1669.

Szolgay, J., Gaal, L., Kohnova, S., Hlavcova, K., Vyleta, R., Bacigal, T., & Blöschl, G. (2015). A process-based analysis of the suitability of copula types for peak-volume flood relationships. In *Proceedings of the International Association of Hydrological Sciences* (pp. 183–188). IAHS. doi:10.5194/piahs-370-183-2015.